Game Theory, Optimization, Reinforcement Learning **NTUA, Spring 2021**

Meeting 1: Dynamical Systems and Game Theory

Advisors: ***Ioannis Panageas, Emmanouil Vlatakis*** **Student:** ***Fivos Kalogiannis***

Objectives

Considering a simple two player game with a utility function $u(x, y)$

- Write down the gradient flow.
- Prove that $(0, 0)$ is a unique Nash Equilibrium.
- Find the distance of (x, y) from $(0, 0)$ if for initial $(x, y) \neq (0, 0)$. Prove that if players x, y were to play based on gradient flow $\text{dist}\left((x, y), (0, 0)\right) = const.$ (η: stepsize for which $0 < \eta < 1$)
- (Discretization) Consider the following rule for updating x, y:

$$\begin{cases} x^{t+1} = x^t - \eta \nabla_x U \\ y^{t+1} = y^t - \eta \nabla_y U \end{cases}$$

i) (Stationary point) Show that if $(x^0, y^0) = (0, 0)$ then $(x^t, y^t) = (0, 0)\ \forall t > 0$.

ii) Use software to simulate the dynamics of the latter game. Pick 3 random initial points (x^0, y^0) for the simulations.

- Prove that distance from $(0, 0)$ per round grows exponentially as a function of stepsize η.

1.1 A simple (continuous) game of two

Consider two player, $x, y \in \mathbb{R}$. They both pick an arbitrary value in the hope of turning a utility function $u : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ to their favor. Player 1, corresponding to x, tries to maximize u while player 2, y, tries to minimize it.

1.1.1 Gradient Flow

We allow every player to move with the same rate towards their goal. Player x will want to maximize u, which means they will move for every dt towards the direction $\nabla_x u$, while player y who seeks to minimize u (conversely, maximize $-u$) will move towards $\nabla_y(-u)$.

For x, we deduce:

$$x' = x - dt \nabla_x u \Rightarrow \tag{1.1}$$

$$x' - x = -dt \nabla_x u \Rightarrow \tag{1.2}$$

$$\frac{x' - x}{dt} = -\nabla_x u \Rightarrow \tag{1.3}$$

$$\frac{dx}{dx} = -\nabla_x u \tag{1.4}$$

Similarly for y, we can write:

$$\frac{dx}{dx} = -\nabla_x(-u) \tag{1.5}$$

Conclusively we define a first order system of differential equations:

$$\begin{cases} \dot{x} = - & \nabla_x u(x,y) \\ \dot{y} = & \nabla_y u(x,y) \end{cases} \Rightarrow \tag{1.6}$$

$$\begin{cases} \dot{x} = - & \nabla_x (xy) \\ \dot{y} = & \nabla_y (xy) \end{cases} \Rightarrow \tag{1.7}$$

$$\begin{cases} \dot{x} = -y \\ \dot{y} = x \end{cases} \tag{1.8}$$

$$\dot{\begin{pmatrix} x \\ y \end{pmatrix}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{1.9}$$

1.1.1.1 Uniqueness of $(0,0)$ as a Nash Equilibrium Point

We remind ourselves that a Nash Equilibrium point is a point (point meaning the tuple of strategy vectors of all players) such that no player can unilaterally alter their strategy to turn utility to their favor if the strategy vectors of all other players were to be fixed.

In our example, the strategy vector of each player is reduced to a scalar real variable. We now move on to prove that only $(x^*, y^*) = (0,0)$ can hold the prerequisites for it to be an equilibrium point.

We observe that if $x = c \neq 0$, player y can alter their strategy to $y \to \infty$ (or $y \to -\infty$) if $c < 0$ (or $c > 0$). The converse holds for $y = c' \neq 0$.

But, if $(x,y) = (0,0)$, we observe that no matter what each player alter their strategy to, utility will stay the same. Hence, $(0,0)$ is a unique equilibrium point for the game in question.

1.1.1.2 Distance from origin

Let us define $V(x,y) = x^2 + y^2$. We observe that V is the squared distance of the point (x,y) from $(0,0)$. We derive with respect to time:

$$\begin{aligned}
\frac{d}{dt}V(x,y) &= \frac{d}{dt}(x^2 + y^2) \Rightarrow \\
\frac{d}{dt}V(x,y) &= 2x\frac{dx}{dt} + 2y\frac{dy}{dt} \Rightarrow \\
\frac{d}{dt}V(x,y) &= 2x(-y) + 2y(x) \Rightarrow \\
\frac{d}{dt}V(x,y) &= 0 \Rightarrow \\
V(x,y) &= const.
\end{aligned}$$

We can thus conclude that the distance of (x,y) remains constant through time.

Alternatively, we can find a closed form solution to 1.9:

$$\begin{cases} x = -c \cdot sin(t) \\ y = c \cdot cos(t) \end{cases} \quad \text{and} \quad \begin{cases} x = c' \cdot sin(t) \\ y = c' \cdot cos(t) \end{cases} \tag{1.10}$$

From either one of the solutions we get the following plot 1.1 of trajectories, either clockwise or anti-clockwise:

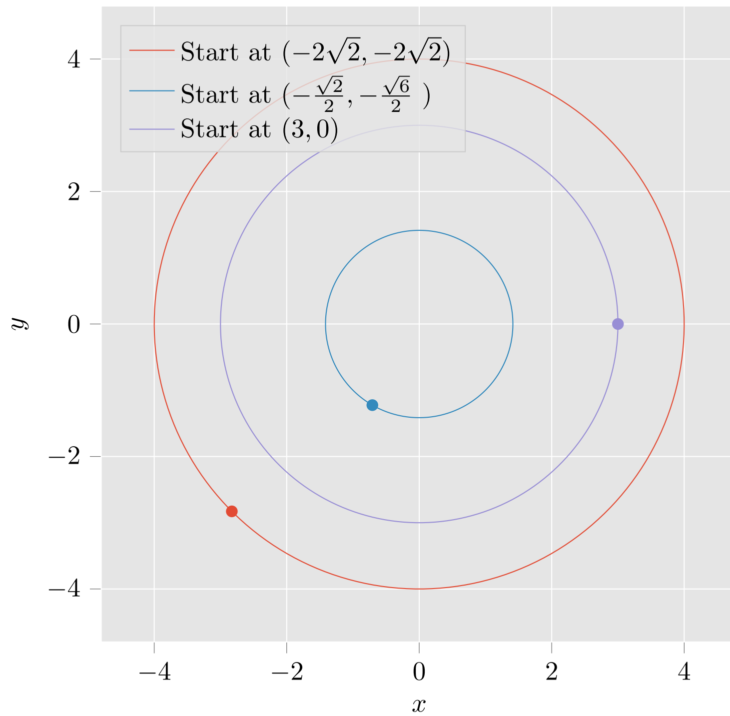

Figure 1.1: Trajectories of (x, y) starting from 3 different initial points

1.1.2 Discretization

Let's us now consider the case of a discrete time version of the latter game. Players still play their moves simultaneously. We will now demonstrate that the vanilla Gradient Descent/Ascent based strategy won't converge. Not only does it not converge, but the distance of (x, y) from $(0, 0)$ will grow exponentially!

Notation

Players are notated with variables x, y, which in our case happen to be scalars but in the general can be vectors. Notation $z^{(n)}$ denotes the value of z at step n. The operator $\nabla_{\boldsymbol{z}}$ stands for the operator of a gradient with respect to $\boldsymbol{z}$ of a given function. (e.g. $\nabla_z(zw) = w$). If $\boldsymbol{z}$ is a vector in $\mathbb{R}^d$, $\nabla_{\boldsymbol{z}}$ is a vector with entries $(\frac{\partial}{\partial z_1}, \dots, \frac{\partial}{\partial z_d})$ and finally η denotes what is called a step size. It merely is a scalar for which $0 < \eta \leq 1$.

1.1.2.1 Vanilla Gradient Descent/Ascent

$$\begin{cases} x^{(n+1)} = x^{(n)} - \eta \nabla_x U \\ y^{(n+1)} = y^{(n)} - \eta \nabla_y U \end{cases} \Rightarrow$$

$$\begin{cases} x^{(n+1)} = x^{(n)} - \eta(-y^{(n)}) \\ y^{(n+1)} = y^{(n)} - \eta(x^{(n)}) \end{cases} \Rightarrow$$

$$\begin{cases} x^{(n+1)} = x^{(n)} + \eta y^{(n)} \\ y^{(n+1)} = y^{(n)} - \eta x^{(n)} \end{cases}$$

Thus we derive the following iterative process:

$$\begin{cases} x^{(n+1)} = x^{(n)} + \eta y^{(n)} \\ y^{(n+1)} = y^{(n)} - \eta x^{(n)} \end{cases} \tag{1.11}$$

1.1.2.2 Non-Convergence of vanilla Gradient Descent/Ascent

We define the ratio of the norm of the (x, y) vector between two sequential iteration of the process as ρ:

$$\rho = \frac{\|(x^{(n+1)}, y^{(n+1)})\|_2}{\|(x^{(n)}, y^{(n)})\|_2}$$

Elaborating more on ρ:

$$\begin{aligned}
&\rho = \frac{\sqrt{(x^{(n+1)})^2 + (y^{(n+1)})^2}}{\sqrt{(x^{(n)})^2 + (y^{(n)})^2}} \Rightarrow \\
&\rho^2 = \frac{(x^{(n+1)})^2 + (y^{(n+1)})^2}{(x^{(n)})^2 + (y^{(n)})^2} \\
&\overset{1.11}{\Longrightarrow} \\
&\rho^2 = \frac{(x^{(n)} + \eta y^{(n)})^2 + (y^{(n)} - \eta x^{(n)})^2}{(x^{(n)})^2 + (y^{(n)})^2} \Rightarrow \\
&\rho^2 = \frac{(x^{(n)})^2 + 2\eta x^{(n)} y^{(n)} + (\eta y^{(n)})^2 + (y^{(n)})^2 - 2y^{(n)} \eta x^{(n)} + (\eta x^{(n)})^2}{(x^{(n)})^2 + (y^{(n)})^2} \Rightarrow \\
&\rho^2 = \frac{(x^{(n)})^2 + \cancel{2\eta x^{(n)} y^{(n)}} + (\eta y^{(n)})^2 + (y^{(n)})^2 - \cancel{2y^{(n)} \eta y^{(n)}} + (\eta x^{(n)})^2}{(x^{(n)})^2 + (y^{(n)})^2} \Rightarrow \\
&\rho^2 = \frac{(x^{(n)})^2 + (\eta y^{(n)})^2 + (y^{(n)})^2 + (\eta x^{(n)})^2}{(x^{(n)})^2 + (y^{(n)})^2} \Rightarrow \\
&\rho^2 = (1 + \eta^2) \frac{(x^{(n)})^2 + (y^{(n)})^2}{(x^{(n)})^2 + (y^{(n)})^2} = (1 + \eta^2) \Rightarrow \\
&\rho = \sqrt{1 + \eta^2}
\end{aligned}$$

It is obvious that $\rho > 0$, more precisely if we define $\rho = 1 + \epsilon$:

$$\begin{aligned}
\rho &= 1 + \epsilon \Rightarrow \\
\sqrt{1 + \eta^2} &= 1 + \epsilon \Rightarrow \\
\epsilon = \sqrt{1 + \eta^2} - 1 &\geq 0 = \mathcal{O}(\eta)
\end{aligned}$$

In the following figure 1.2 we can observe how will the vector (x, y) move trough time if it were to start from 3 different random initial points.

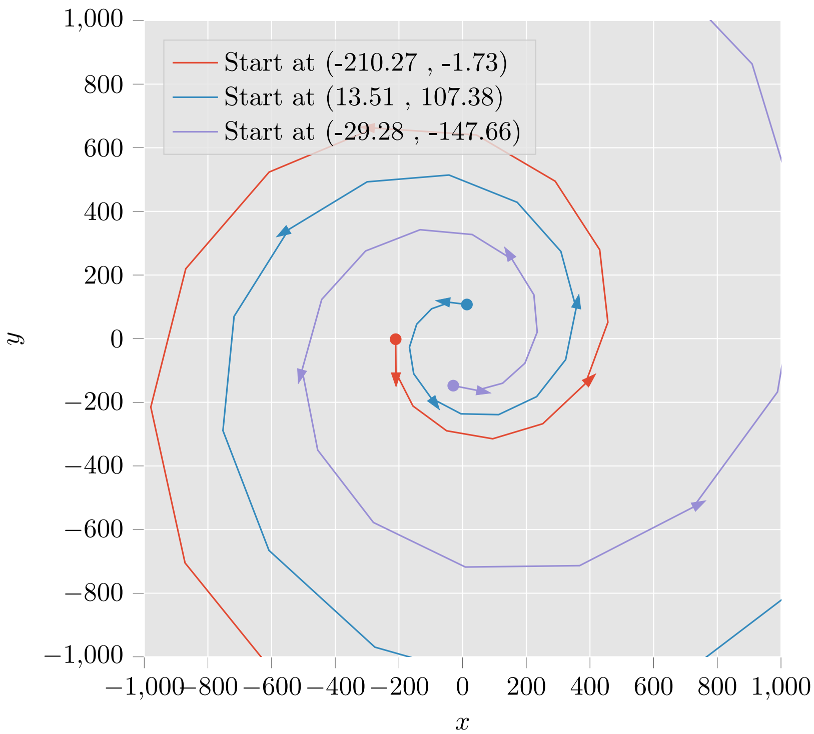

Figure 1.2: Trajectories of (x, y) starting from 3 different initial points

References

BOYCE, W. E., DIPRIMA, R. C. and MEADE, D. B. (2017). *Elementary differential equations*. John Wiley & Sons.